Galois representations attached to $\mathbb{Q}$-curves and the generalized Fermat equation $A^4 + B^2 = C^p$

Jordan S. Ellenberg *
Princeton University
ellenber@math.princeton.edu

22 Jul 2003

Abstract

We prove that the equation $A^4+B^2 = C^p$ has no solutions in coprime positive integers when $p \geq 211$. The main step is to show that, for all sufficiently large primes p, every $\mathbb{Q}$-curve over an imaginary quadratic field K with a prime of bad reduction greater than 6 has a surjective mod p Galois representation. The bound on p depends on K and the degree of the isogeny between E and its Galois conjugate, but is independent of the choice of E. The proof of this theorem combines geometric arguments due to Mazur, Momose, Darmon, and Merel with an analytic estimate of the average special values of certain L-functions.

1 Introduction

The resolution of the Fermat problem has demonstrated a close relationship between the solutions of Diophantine equations and the arithmetic of abelian varieties over number fields. It remains far from clear which Diophantine equations can be productively studied along the lines developed by Frey, Hellegouarch, Serre, Ribet, Wiles, and Taylor.

In particular, one wonders whether modular abelian varieties can address the classical problem of describing all solutions to the *generalized Fermat equation*

$$A^p + B^q = C^r \tag{1.1}$$

in coprime integers A, B, C. Darmon and Granville [10] have proved that (1.1) has only finitely many solutions for any particular p, q, r satisfying $1/p + 1/q + 1/r < 1$. It is conjectured that (1.1) has only finitely many solutions, excepting $1^p + 2^3 = 3^2$, even if p, q, r are allowed to vary (still subject to $1/p + 1/q + 1/r < 1$.)

Certain special cases of (1.1) and similar equations have been treated by Darmon, Merel, and Ribet([7],[11],[26]) using elliptic curves over $\mathbb{Q}$, and by Bruin [4] using Chabauty methods. In [9], Darmon discusses the relationship between more general cases of (1.1) and as-yet-unproved conjectures about the Galois representations attached to Hilbert-Blumenthal abelian varieties over number fields.

Our goal in the present paper is twofold. Our main motivation (or, as Darmon and Merel put it in [11], our "excuse") is to prove the following Diophantine theorem:

*Partially supported by NSA Young Investigator Grant MDA905-02-1-0097.

Theorem (Theorem 4.1**).** *Suppose A, B, C are coprime integers such that*

$$A^4 + B^2 = C^p \tag{1.2}$$

and $p \geq 211$. Then $AB = 0$.

We will attach an elliptic curve E to any solution to (1.2). However, E will be defined not over $\mathbb{Q}$, but over $\mathbb{Q}[i]$, and it will be isogenous to its Galois conjugate. An elliptic curve, like E, whose isogeny class is defined over $\mathbb{Q}$ is called a $\mathbb{Q}$-*curve*. (The idea of studying (1.2) by means of $\mathbb{Q}$-curves was arrived at independently by Darmon in [8].)

In order to prove Theorem 4.1, it is necessary to bring our knowledge of the arithmetic of $\mathbb{Q}$-curves more in line with our knowledge about elliptic curves defined over $\mathbb{Q}$. Our second goal in this paper is to prove a theorem on surjectivity of mod p representations attached to $\mathbb{Q}$-curves with non-integral j-invariant. If K is a quadratic number field, a $\mathbb{Q}$-curve E/K of degree d is an elliptic curve over K which admits a cyclic isogeny of degree d to its Galois conjugate.

Theorem (Theorem 3.14). *Let K be an imaginary quadratic field and d a square-free positive integer. There exists an effective constant $M_{K,d}$ such that, for all primes $p > M_{K,d}$ and all $\mathbb{Q}$-curves E/K of degree d, either*

- *the representation*
$$\mathbb{P}\bar{\rho}_{E,p} : \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{PGL}_2(\mathbb{F}_p)$$
is surjective, or
- *E has potentially good reduction at all primes not dividing* 6*.*

Theorem 3.14 plays the role that Mazur's theorem [23] does in the solution of Fermat's problem. It is worth remarking that the analogue of Theorem 3.14 for elliptic curves over $\mathbb{Q}$ is still conjectural.

We also need a modularity theorem for $\mathbb{Q}$-curves; we have proved the result we need in an earlier paper with C. Skinner [15].

Much of the proof of Theorem 3.14 follows along the lines of work of Mazur [23] and Momose [25] for elliptic curves over $\mathbb{Q}$. However, at a key point we must introduce an analytic argument on average special values of L-functions; it is because of the analytic step that we can prove the main theorem of the paper only for $p \geq 211$. It should be emphasized that the remaining cases can be handled by a finite, though at present unfeasible, computation. In particular, the methods of this paper yield the following fact (Proposition 4.7)

Theorem. *Let $p > 13$ be a prime. Suppose there exists either*

- *a newform in $S_2(\Gamma_0(2p^2))$ with $w_p f = f$ and $w_2 f = -f$; or*
- *a newform in $S_2(\Gamma_0(p^2))$ with $w_p f = f$,*

such that $L(f \otimes \chi, 1) \neq 0$, where χ is the Dirichlet character of conductor 4*. Then the equation $A^4 + B^2 = C^p$ has no primitive non-trivial solutions.*

The author is grateful to Henri Darmon and Emmanuel Kowalski for useful discussions about the theorems proved here, and to the referee for thorough and helpful remarks.

2 Twisted modular curves and their Jacobians

A main theme of the present article is the analysis of certain twisted versions of modular curves. We begin with the "untwisted" versions. Recall that, for any prime p, the curve $X^{split}(p)$ is a coarse moduli space parametrizing pairs $(E, \{A, B\})$ where E is an elliptic curve and $\{A, B\}$ is an unordered pair of distinct cyclic subgroups of $E[p]$. Similarly, $X^{ns}(p)$ parametrizes pairs (E, N) where N is a pair of Galois-conjugate points in $\mathbb{P}E[p] \otimes_{\mathbb{F}_p} \mathbb{F}_{p^2}$. We call such an N a *non-split structure* on E. Note that $X^{split}(p)$ and $X^{ns}(p)$ are known to have smooth models over $\mathbb{Z}[1/p]$.

Definition 2.1. Let p be a prime and m a positive integer prime to p. We define

$$X_0^s(m;p) = X_0(m) \times_{X(1)} X^{split}(p)$$

and

$$X_0^{ns}(m;p) = X_0(m) \times_{X(1)} X^{ns}(p).$$

The curves $X_0(mp), X_0^s(m;p)$ and $X_0^{ns}(m;p)$ have involutory automorphisms w_m arising from the Fricke involution w_m on $X_0(m)$. Let K be a quadratic field. For $X = X_0(mp), X_0^s(m;p)$ or $X_0^{ns}(m;p)$, let $X^K/\mathbb{Q}$ be the twisted form of X admitting an isomorphism $\phi : X^K/K \to X/K$ such that $\phi^\sigma = w_m \circ \phi$. Note that $X^K(\mathbb{Q})$ can be described as the subset of $P \in X(K)$ such that $P^\sigma = w_m P$, for σ a generator of $\mathrm{Gal}(K/Q)$.

Recall that a $\mathbb{Q}$-*curve* E/K is an elliptic curve which is isogenous to its Galois conjugate. If there exists such an isogeny of degree d, we say E/K is a $\mathbb{Q}$-curve of degree d.

In [15, Prop. 2.3] we define a projective mod p Galois representation

$$\mathbb{P}\bar{\rho}_{E,p} : \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{PGL}_2(\mathbb{F}_p)$$

associated to any $\mathbb{Q}$-curve.

We will later show that under certain circumstances, $\mathbb{P}\bar{\rho}_{E,p}$ has large image. Our main tool is the following proposition.

Proposition 2.2. *Let d be a square-free positive integer, and let E/K be a $\mathbb{Q}$-curve of degree d over a quadratic number field K. Let p be a prime not dividing d.*

Suppose the image of $\mathbb{P}\bar{\rho}_{E,p}$ lies in a Borel subgroup of $\mathrm{PGL}_2(\mathbb{F}_p)$ (resp. the normalizer of a split Cartan subgroup, normalizer of a non-split Cartan subgroup.) Then the point of $X_0(dp)(K)$ (resp. $X_0^s(d;p)(K)$, $X_0^{ns}(d;p)(K)$) corresponding to E is a point of $X_0(dp)^K(\mathbb{Q})$ (resp. $X_0^s(d;p)^K(\mathbb{Q})$, $X_0^{ns}(d;p)^K(\mathbb{Q})$.

Proof. We will discuss the case of $X_0(dp)$; the other two are similar. Let C_p be the cyclic subgroup of $E[p]/K$ which is fixed by $\mathrm{Gal}(\bar{K}/K)$. Let $\mu : E^\sigma \to E$ be the degree d isogeny, and let $C_d = \mu(E^\sigma[d])$. Now $\mathbb{P}\rho_{E,p}(\sigma)$ acts on $\mathbb{P}E[p]$ by sending x to $\mu(x^\sigma)$. Since, by hypothesis, $\mathbb{P}\rho_{E,p}(\sigma)$ fixes C_p, we have that $C_p^\sigma = \mu^{-1}C_p$. So indeed

$$(E, C_d, C_p)^\sigma = (E/C_d, E[d]/C_d, \mu^{-1}C_p) = w_d(E, C_d, C_p),$$

which proves the desired result. □

3 Good reduction for $\mathbb{Q}$-curves with rational level structures

In this section we use arguments derived from Mazur's foundational paper [23] in order to show that points on twisted modular curves of large level must have good reduction at all large primes.

Proposition 3.1. *Let R be a finite extension of $\mathbb{Z}_\ell$ with fraction field L and maximal ideal λ, and let X/R be a stable curve. Write X^{smooth} for the smooth part of X. Suppose A/R is a semi-abelian scheme with a morphism $\phi : X^{smooth} \to A$. Let x and y be distinct sections in $X^{smooth}(R)$, such that $\phi(x) = 0$. Write x_λ and y_λ for the restrictions of x and y to the closed fiber of* $\operatorname{Spec} R$.

Then suppose

- *The absolute ramification index e_ℓ of R is less than $\ell - 1$;*
- *ϕ is a formal immersion at x_λ.*
- *The restriction of $\phi(y) - \phi(x)$ to $A(L)$ is of finite order.*

Then x_λ and y_λ are distinct points of $X^{smooth}(R/\lambda)$.

Proof. The proposition is essentially Corollary 4.3 of [23]. For the reader's convenience, we recount the argument here. Let $q \in A(R)$ be the point $\phi(y) - \phi(x)$.

First of all, suppose that the restriction to q to $A(L)$ is a non-zero torsion point of order m. Then it follows from a theorem of Raynaud [23, Prop. 1.1] that the specialization of q to the closed point of $\operatorname{Spec} R$ also has exact order m. In particular, q does not reduce to 0 mod λ. Thus, y does not reduce to x mod λ.

So we may assume that $\phi(y) = \phi(x)$. Suppose $x_\lambda = y_\lambda$. The section

$$x : \operatorname{Spec} R \to X$$

restricts to a map from $\operatorname{Spec} R$ to the spectrum of the completed local ring $\hat{\mathcal{O}}_{X,x_\lambda}$. So x yields a map of rings

$$\tilde{x} : \hat{\mathcal{O}}_{X,x_\lambda} \to R.$$

Likewise, y yields a map

$$\tilde{y} : \hat{\mathcal{O}}_{X,y_\lambda} = \hat{\mathcal{O}}_{X,x_\lambda} \to R.$$

Since $x \neq y$, the two maps $\tilde{x}$ and $\tilde{y}$ are distinct. Let 0_λ be the identity in $A(R/\lambda)$. Then the fact that ϕ is a formal immersion along x means precisely that the map of completed local rings

$$\tilde{\phi} : \hat{\mathcal{O}}_{A,0_\lambda} \to \hat{\mathcal{O}}_{X,x_\lambda}$$

is a surjection. But this contradicts the fact that $\tilde{x} \circ \tilde{\phi} = \tilde{y} \circ \tilde{\phi}$. □

Proposition 3.2. *Let K be a quadratic field, and E/K be a $\mathbb{Q}$-curve of squarefree degree d. Suppose $\mathbb{P}\bar{\rho}_{E,p}$ is reducible for some $p = 11$ or $p > 13$ with $(p, d) = 1$. Then E has potentially good reduction at all primes of K of characteristic greater than 3.*

Proof. Let $C/\mathbb{Q}$ be the twisted modular curve $X_0(dp)^K$, and take $J = \operatorname{Jac}(C)$. By Proposition 2.2, E corresponds to a point $P \in C(\mathbb{Q})$. We will think of P as lying in $C(K) = X_0(dp)(K)$. Let λ be a prime of K of characteristic greater than 3. We denote by ∞ the usual cusp of $X_0(dp)$, and by ∞_p the corresponding cusp of $X_0(p)$.

Define a map $g : X_0(dp) \to J_0(p)$ by the rule

$$g((E, C_d, C_p)) = [E, C_p] + [E/C_d, C_p/C_d] - 2[\infty_p].$$

Let f be a Hecke eigenform in $S_2(\Gamma_0(p))$. The map $\pi_f \circ g : X_0(dp) \to A_f$ extends by the Néron mapping property to a map, also denoted $\pi_f \circ g$, from the smooth part of $X^0(dp)/\mathbb{Z}$ to the Néron model of A_f, which we also denote A_f.

Lemma 3.3. *The map $\pi_f \circ g : X_0^{smooth}(dp)/\mathbb{Z}[1/2] \to A_f/\mathbb{Z}[1/2]$ is a formal immersion at the point $\infty \in X_0^{smooth}(dp)(\mathbb{F}_\lambda)$.*

Proof. First of all, $\pi_f \circ g$ induces an isomorphism of the residue fields attached to the points $\infty \in X_0^{smooth}(dp)(\mathbb{F}_\lambda)$ and $0 \in A_f(\mathbb{F}_\lambda)$, both of which are $\mathbb{F}_\lambda$. So it suffices to show that the map

$$(\pi_f \circ g)^* : \mathrm{Cot}_0(A_f/\mathbb{F}_\lambda) \to \mathrm{Cot}_\infty(X_0)(dp)/\mathbb{F}_\lambda$$

is a surjection [17, 17.4.4].

Let $g_1 : X_0^{smooth}(dp)/\mathbb{Z}[1/2] \to J_0(p)/\mathbb{Z}[1/2]$ be the morphism sending the point (E, C_d, C_p) to the divisor $[(E, C_p)]-[\infty]$. Likewise, let g_d be the morphism sending (E, C_d, C_p) to $[(E/C_d, C_p/C_d)]-[\infty]$. Then $g = g_1 + g_d$. Now $\pi_f \circ g_1$ factors as

$$X_0(dp) \xrightarrow{p} X_0(p) \xrightarrow{j} A_f,$$

where p, the "forgetting of d-structure" morphism, is unramified at ∞ and j is a formal immersion at the point at infinity of $X_0(p)(\mathbb{F}_\lambda)$ ([23, Prop 3.1].) It follows that the composition $\pi_f \circ g_1$ is a formal immersion at ∞, whence $(\pi_f \circ g_1)^*$ is a surjection on cotangent spaces. On the other hand, $\pi_f \circ g_d$ factors as

$$X_0(dp) \xrightarrow{p \circ w_d} X_0(p) \xrightarrow{j} A_f$$

But $p \circ w_d$ is ramified at ∞, so the image of the map $(p \circ w_d)^*$ in $\mathrm{Cot}_\infty(X_0)(dp)/\mathbb{F}_\lambda$ is zero. Therefore, $(\pi_f \circ g_d)^*$ is also the zero map on cotangent spaces. So $(\pi_f \circ g)^* = (\pi_f \circ g_1)^*$ is surjective, as desired. □

By the hypotheses on p, we can, and do, choose an eigenform $f \in S_2(\Gamma_0(p))$ such that $A_f(\mathbb{Q})$ is a finite group. For instance, we may choose f so that A_f lies in the Eisenstein quotient of $J_0(p)$. We now want to derive a contradiction from 3.1.

Suppose E has potentially multiplicative reduction at λ. Applying an Atkin-Lehner involution if necessary, we may assume that P reduces to ∞ mod λ.

We will now apply Proposition 3.1, where $L = K_\lambda, X = X_0(dp)$, and $A = A_f$. Take x to be the section ∞, and y to be the section given by P. It follows from the potentially multiplicative reduction of E that y is a section of $X_0(dp)^{smooth}/\mathcal{O}_{K_\lambda}$.

Finally, take ϕ to be the map $\pi_f \circ g$. Since L is a quadratic extension of $\mathbb{Q}_\ell$, and $\ell > 3$, we have $e_\ell < \ell - 1$. By Lemma 3.3, ϕ is a formal immersion at $x_\lambda = \infty_\lambda$.

Let σ be the nontrivial element of $\mathrm{Gal}(K/\mathbb{Q})$; then

$$g(P)^\sigma = [p(P^\sigma)] + [p(w_d P^\sigma)] - 2[\infty_p] = [p(w_d P)] + [p(P)] - 2[\infty_p] = g(P).$$

In particular, $\phi(y) = \pi_f(g(P))$ lies in $A_f(\mathbb{Q})$, and therefore has finite order. Note also that $\phi(x) = \pi_f(g(\infty)) = 0$.

We can conclude that $x_\lambda \neq y_\lambda$, contradicting the hypothesis that P reduces to ∞ mod λ. □

Proposition 3.4. *Let K be a quadratic field, and E/K be a $\mathbb{Q}$-curve of squarefree degree d. Suppose the image of $\mathbb{P}\bar{\rho}_{E,p}$ lies in the normalizer of a split Cartan subgroup of $\mathrm{PGL}_2(\mathbb{F}_p)$, for $p = 11$ or $p > 13$ with $(p, d) = 1$. Then E has good reduction at all primes of K not dividing 6.*

Proof. This case is very similar to that of Proposition 3.2. Here, the key lemma on formal immersions is due to Momose [25].

Let C be the twisted modular curve $X_0^s(d;p)^K$. Then E yields a point

$$P \in C(\mathbb{Q}) \subset C(K) = X_0^s(d;p)(K).$$

To be more precise, $P \in C(K)$ parametrizes the triple $(E, C_d^0, \{A_p^0, B_p^0\})$ where C_d^0 is the kernel of the isogeny between E and its Galois conjugate, and $\{A_p^0, B_p^0\}$ is the chosen pair of cyclic subgroups fixed by the action of $\mathrm{Gal}(\bar{K}/K)$.

Let $X_0^{sc}(d;p)$ be the modular curve parametrizing quadruples (E, C_d, A_p, B_p) where C_d is a cyclic subgroup of E of order d and A_p and B_p are linearly independent cyclic subgroups of E of order p. Similarly, let $X^{sc}(p)$ be the curve parametrizing triples (E, A_p, B_p). Note that $X_0^{sc}(d;p)$ (resp. $X^{sc}(p)$) is naturally a double cover of $X_0^s(d;p)$ (resp. $X^{split}(p)$). Write w_p for the involution of $X_0^{sc}(d;p)$ switching A_p and B_p. A cusp of $X_0^{sc}(d;p)$ is determined by its image in $X_0(d)$ and $X^{sc}(p)$. We write (c, c') for the cusp of $X_0^{sc}(d;p)$ whose image in $X_0(d)$ is c and whose image in $X^{sc}(p)$ is c'.

Let

$$h : X_0^{sc}(d;p)/\mathbb{Q} \to J_0(p)/\mathbb{Q}$$

be the map defined by

$$h((E, C_d, A_p, B_p)) = [(E, A_p)] - [(E/B_p, E[p]/B_p)] + [(E/C_d, A_p/C_d)] - [(E/(B_p+C_d), E[p]/(B_p+C_d))].$$

Then we have a diagram of schemes over $\mathbb{Q}$

$$\begin{array}{ccc} X_0^{sc}(d;p) & \longrightarrow & X_0^s(d;p) \\ {\scriptstyle h}\big\downarrow & & {\scriptstyle h^-}\big\downarrow \\ J_0(p) & \longrightarrow & (1-w_p)J_0(p) \end{array}$$

where h^- is defined to make the above diagram commute.

The moduli problem over $\mathrm{Spec}\,\mathbb{Q}$ coarsely represented by $X_0^{sc}(d;p)$ can be extended to a moduli problem over $\mathrm{Spec}\,\mathbb{Z}$ as in [19]; this moduli problem is coarsely represented by a curve over $\mathrm{Spec}\,\mathbb{Z}$, which we also denote $X_0^{sc}(d;p)$. (Note that $X_0^{sc}(d;p)$ is isomorphic to $X_0(dp^2)$.) The curve $X_0^{sc}(d;p)$ is smooth away from characteristics dividing dp. In characteristics dividing d, the reduction of $X_0^{sc}(d;p)$ is smooth away from supersingular points, and in particular is smooth at all cusps. In characteristic p, the reduction of $X_0^{sc}(d;p)$ is made up of three components, parametrizing triples (E, C_d, A_p, B_p) where, respectively:

- $A_p \cong \mu_p$ étale-locally;
- $B_p \cong \mu_p$ étale-locally;
- neither A_p nor B_p is étale-locally isomorphic to μ_p.

The smooth part of the special fiber at p contains the ordinary locus in the first two components. (See [25, §1] and [19, (13.5.6)] for facts used here about the special fiber.)

By the Néron mapping property, h extends to a map from $X_0^{sc;smooth}(d;p)/\mathbb{Z}$ to the Néron model of $J_0(p)$.

Let f be a Hecke eigenform in $S_2(\Gamma_0(p))$ such that $w_p f = -f$. Then we have a projection map from $J_0(p)/\mathbb{Z}$ to $A_f/\mathbb{Z}$.

We base change this map to $\mathbb{Z}[\zeta_p][1/2]$ to obtain a map

$$\pi_f : J_0(p)/\mathbb{Z}[\zeta_p][1/2] \to A_f/\mathbb{Z}[\zeta_p][1/2].$$

Lemma 3.5. *Let λ be a prime of $\mathbb{Q}(\zeta_p)$ with residue field $\mathbb{F}_\lambda$. The composition $\pi_f \circ h$ is a formal immersion at every cusp (∞, c') of $X_0^{sc;smooth}(d;p)_{\mathbb{F}_\lambda}$, for all $\lambda \not| 2p$. For all $\lambda \not| 2$, the composition $\pi_f \circ h$ is a formal immersion at the cusp (∞, ∞) of $X_0^{sc;smooth}(d;p)_{\mathbb{F}_\lambda}$.*

Proof. We proceed much as we did in Lemma 3.3. First, note that the hypotheses on ℓ guarantee that (c, c') lies in the smooth locus of $X_0^{sc}(d;p)$ as in [25, §1].

Each cusp of $X_0^{sc}(d;p)_{\mathbb{F}_\lambda}$ is defined over $\mathbb{F}_\lambda$ ([25, proof of (2.5)]), and $\pi_f \circ h$ is defined over $\mathbb{F}_\lambda$ by definition; it follows that for each cusp c, the residue field of c and the residue field of $\pi_f \circ h(c)$ are both $\mathbb{F}_\lambda$. So in order to prove the lemma, it suffices to show that the map

$$(\pi_f \circ h)^* : \mathrm{Cot}_0(A_f/\mathbb{F}_\lambda) \to \mathrm{Cot}_{(\infty,c')} X_0^{sc}(d;p)/\mathbb{F}_\lambda$$

is a surjection.

We now write $h = h_1 + h_d$, where

$$h_1((E, C_d, A_p, B_p)) = [(E, A_p)] - [(E/B_p, E[p]/B_p)]$$

and $h_d = h_1 \circ w_d$. Now $\pi_f \circ h_1$ factors as

$$X_0^{sc;smooth}(d;p) \stackrel{a}{\to} X^{sc;smooth}(p) \stackrel{\pi_f \circ g}{\to} A_f.$$

Here g is the morphism from $X^{sc;smooth}(p)$ to $J_0(p)$ defined by Momose [25, §2], and a is the "forgetting of d-structure" morphism. Again, a is unramified at (∞, c'), and $\pi_f \circ g$ is a formal immersion at the cusp c' of $X^{sc;smooth}(p)$, by [25, (2.5)]. So $(\pi_f \circ h_1)^*$ is a surjection on cotangent spaces.

On the other hand, $w_d \circ a$ is ramified at (∞, c'), so $(\pi_f \circ h_d)^*$ is the zero map on cotangent spaces. So $(\pi_f \circ h)^*$ is surjective, as desired. □

We now make the further stipulation on f that $A_f(\mathbb{Q})$ is a finite group. (Again, we may choose f such that A_f is a quotient of the Eisenstein quotient of $J_0(p)$.) Let σ be the nontrivial element of $\mathrm{Gal}(K/\mathbb{Q})$. We have

$$(E, C_d^0, \{A_p^0, B_p^0\})^\sigma = (E/C_d^0, E[d]/C_d^0, \{A_p^0/C_d^0, B_p^0/C_d^0\}).$$

It follows immediately that $h^-(P^\sigma) = h^-(P)$. Let $Q \in X_0^{sc}(d;p)(\bar{\mathbb{Q}})$ be a point lying over P. Then $h(Q) = h^-(P) \in [(1 - w_p)J_0(p)](\mathbb{Q})$, so $\pi_f(h(Q))$ lies in $A_f(\mathbb{Q})$, and is thus of finite order.

Let M be the field of definition of Q. Let G be the group of automorphisms of $X_0^{sc}(d;p)$ generated by w_d and w_p; then $G \cong (\mathbb{Z}/2\mathbb{Z})^2$ and Q lies over a $\mathbb{Q}$-point of $X_0^{sc}(d;p)/G$, whence M is a subfield of a biquadratic field over $\mathbb{Q}$. Suppose λ is a prime of M such that E has potentially multiplicative reduction at λ. Write $(c, c') \in X_0^{sc}(d;p)(\bar{\mathbb{Q}})$ for the cusp to which Q reduces mod λ. Applying w_d if necessary, we may assume that $c = \infty$.

If $\lambda \not| p$, the map $\pi_f \circ h$ is a formal immersion at $(c, c')_\lambda$ by Lemma 3.5. Suppose $\lambda | p$. Since E has potentially multiplicative reduction, it acquires semistable reduction after a quadratic extension M' of M; since M is biquadratic, the absolute ramification index of M' at p is at most 4. If Q reduces to a cusp other than 0 and ∞, the group schemes A_p^0 and B_p^0 are both étale over $\mathcal{O}_{M'}$ ([25, proof of (2.5)] whence, by Weil pairing, μ_p is étale over $\mathcal{O}_{M'}$; this makes the absolute ramification index of M' over p at least $p-1$, a contradiction. If Q reduces to 0, we can act on Q by w_p to make $c' = \infty$. Now, by Lemma 3.5, the map $\pi_f \circ h$ is a formal immersion at $(c, c')_\lambda$.

We will now apply Proposition 3.1, using $L = M_\lambda, X = X_0^{sc}(d;p)$, and $A = A_f$. Take x to be the cuspidal section (c, c'), and y to be the section given by Q. Finally, take ϕ to be the map $\pi_f \circ h$.

Since M is a subfield of a biquadratic extension of $\mathbb{Q}$, its ramification degree is at most 2 over any odd prime. So we have $e_\ell < \ell - 1$. Now the conclusion of Proposition 3.1 contradicts the hypothesis that Q and (c, c') reduce to the same point of $X_0^{sc}(d;p)(M/\lambda)$. □

We now turn to the case of $\mathbb{Q}$-curves E whose mod p Galois representations have image in the normalizer of a non-split Cartan subgroup. This case is more difficult, due to the absence of rank 0 quotients of $J^{ns}(p)$. However, we show below by analytic means that the Jacobian of the twisted modular curve $X_0^{ns}(d;p)^K$ does have rank 0 quotients; we then obtain a good reduction theorem on E using a formal immersion result of Darmon and Merel [11].

Proposition 3.6. *Let K be an imaginary quadratic field, and E/K be a $\mathbb{Q}$-curve of squarefree degree d. There exists a constant $M_{d,K}$ with the following property.*

Suppose the image of $\mathbb{P}\bar{\rho}_{E,p}$ lies in the normalizer of a non-split Cartan subgroup of $\mathrm{PGL}_2(\mathbb{F}_p)$, for $p > M_{d,K}$. Then E has potentially good reduction at all primes of K.

Remark 3.7. All is not lost if K is a real quadratic field. We will see below that when K is imaginary, there exists a newform f on level p^2 satisfying the conditions of Proposition 3.9. When K is real, there still may be newforms f on other levels satisfying those conditions. However, the methods described here cannot treat the case of arbitrary K and d. For instance, suppose d is prime, and K is a real quadratic field in which d is inert. Then if f is either a newform in $S_2(\Gamma_0(p^2))$ with $w_p f = f$, or a newform in $S_2(\Gamma_0(dp^2))$ with $w_p f = f$ and $w_d f = -f$, we have by a theorem of Weil [21, Th. 6] that $f \otimes \chi_K$ has negative functional equation, so that $L(f \otimes \chi_K, 1) = 0$. This leaves us in much the same position as one who tries to use Mazur's method to control points on $X^{ns}(p)$; the Jacobian of the curve in question, assuming Birch-Swinnerton-Dyer, has no rank 0 quotient.

Proof. Let C be the twisted modular curve $X_0^{ns}(d;p)^K$, and write $J = \mathrm{Jac}(C)$. Then E yields a point $P \in C(\mathbb{Q})$.

We have an isomorphism $J \times_{\mathbb{Q}} K \cong J_0^{ns}(d;p)/K$; from a result of Chen and Edixhoven [6],[12] there is a surjective homomorphism

$$\alpha : J_0^{ns}(d;p) \to J_0'(dp^2)/w_p.$$

where $J_0'(dp^2)$ is the p-new quotient of $J_0(dp^2)$.

Let $\mathbf{T}$ be the algebra generated by Hecke operators of degrees prime to dp, together with the group W of Atkin-Lehner involutions of degrees dividing d. It follows from Theorem 2 of [12] that the map α is compatible with the action of $\mathbf{T}$ on either side (in that theorem, let $\mathcal{C}$ be the isogeny category of abelian varieties endowed with an action of $\mathbf{T}$, and M the Jacobian of $X_0(d) \times_{X(1)} X(p)$.)

Suppose f is either

- a newform in $S_2(\Gamma_0(dp^2))$ with $w_p f = f$ and $w_d f = -f$;
- a newform in $S_2(\Gamma_0(d'p^2))$ with d' a proper divisor of d and $w_p f = f$.

In each case, we have a quotient morphism

$$\pi_f : J_0(dp^2) \to A_f$$

such that the action of w_d on $J_0(dp^2)$ induces the involution -1 on A_f. In case d is not prime, we can and do choose our π_f such that the quotient A_f is preserved by the whole group W of Atkin-Lehner

involutions. More precisely: for each $e|(d/d')$, we have a map $B_e : J_0(d'p^2) \to J_0(dp^2)$. Choose a character $\chi : (\mathbb{Z}/(d/d')\mathbb{Z})^* \to \pm 1$. It follows from Lemma 26 of [3] that the quotient map

$$I_\chi : \sum_{e|(d/d')} eB_e\chi(e) : J_0(d'p^2) \to J_0(dp^2)$$

has image stable under the action of W, and on which the action of w_d on the quotient is $w_{d'}$ twisted by the scalar $\chi(d)$. So if we choose χ such that $\chi(d)$ and the eigenvalue of $w_{d'}$ on A_f have opposite signs, then the image under I_χ of $A_f \subset J_0(d'p^2)$ is a subvariety of $J_0(dp^2)$, isogenous to A_f, which is stable under W and on which w_d acts as -1. Now let π_f be projection onto that subvariety.

Composing with α yields a morphism from $J_0^{ns}(d;p)$ to A_f; replacing A_f by an isogenous variety A'_f, we have a quotient

$$\pi'_f : J_0^{ns}(d;p) \to A'_f$$

which is compatible with the action of $\mathbf{T}$ and has connected kernel. In particular, $\pi'_f \circ w_d = -\pi'_f$.

Denote χ_K by χ, and write $A_f \otimes \chi$ for the twist of A_f by χ. Let σ be the non-trivial element of $\mathrm{Gal}(K/\mathbb{Q})$. Then we have a commutative diagram of abelian varieties over K:

$$\begin{array}{ccc} J_0^{ns}(d;p) & \xrightarrow{\pi'_f} & A'_f \\ {\scriptstyle i}\downarrow & & {\scriptstyle j}\downarrow \\ J & \longrightarrow & A'_f \otimes \chi \end{array}$$

where i and j are isomorphisms such that $i^\sigma = w_d \circ i$ and $j^\sigma = -j$. Let

$$\psi_f : J \to A'_f \otimes \chi$$

be the composition $j \circ \pi'_f \circ i^{-1}$. Then

$$\psi_f^\sigma = j^\sigma \circ (\pi'_f)^\sigma \circ (i^{-1})^\sigma = (-j) \circ \pi'_f \circ w_d \circ i^{-1} = (-j) \circ (-\pi'_f) \circ i^{-1} = \psi_f.$$

In other words, ψ_f is defined over $\mathbb{Q}$.

Let R_0 be the ring of integers of the number field $K(\zeta_p + \zeta_p^{-1})$, and let $R = R_0[1/2dp]$. Then $X_0^{ns}(d;p)$ has a smooth model over R and the cusp ∞ of $X_0^{ns}(d;p)$ is defined over R [11, §5].

We can define a map

$$h : X_0^{ns}(d;p)/R \to J_0^{ns}(d;p)/R$$

by setting $h(P) = [P] - [\infty]$.

Lemma 3.8. *Let λ be a prime of R. Then the map*

$$\pi'_f \circ h : X_0^{ns}(d;p)/R \to A'_f/R$$

is a formal immersion at the point $\bar{\infty}$ of $X_0^{ns}(d;p)(\mathbb{F}_\lambda)$.

Proof. This fact is almost precisely Lemma 8.2 of [11].

One difference is that our quotient A'_f is not preserved by T_n for all n prime to p, but by T_n for all n prime to dp and all $w_{d'}$ for $d'|d$. We need to prove that, as in [11], there exists a differential form ω on A'_f whose associated modular form $\tilde{g} = \sum a_n(\tilde{g})q^{n/p}$ has $a_1(\tilde{g}) \neq 0 \pmod{\lambda}$. In fact, we will prove this for any quotient A of $J_0^{ns}(d;p)$ which is preserved by $\mathbf{T}$ and which is not killed by α.

Write $S(A)$ for the vector space of weight 2 cusp forms attached to A. Suppose that $a_1(\tilde{g}) = 0 \pmod{\lambda}$ for every $\tilde{g}$ in $S(A)$.

Choose some $\tilde{g}$ in $S(A)$ which is an eigenform for $\mathbf{T}$, which is not in the kernel of α, and which does not reduce to 0 mod λ. The form

$$g = \sum a_n(\tilde{g}) q^n$$

is a form on $\Gamma_1(dp^2)$ which is also an eigenform for $\mathbf{T}$. (We remark, however, that the Hecke eigenform on $\Gamma_0(dp^2)$ associated to $\tilde{g}$ via the map α does not necessarily have the same Hecke eigenvalues as g.)

Let g_0 be a newform on some level dp^2/M with the same eigenvalues as g. Then we can write

$$g = \sum_{d'|M} \alpha_{d'} B_{d'} g_0.$$

where B_d is the Hecke operator sending $f(\tau)$ to $f(d\tau)$. (In this and all other discussion of Hecke operators, we follow the notation of [3].)

Suppose $\alpha_e \neq 0 \pmod{\lambda}$ for some $e|M$ with $(e, p) = 1$. By [2, Prop. 1.5],

$$w_e g = c\alpha_e B_1 g_0 + \sum_{d'|M, d'>1} B_{d'} h_{d'}$$

where c is a constant not divisible by λ and the $h_{d'}$ are other modular forms. Since $a_1(B_{d'} h_{d'}) = 0$ for all $d' > 1$, we see that $a_1(w_e g) \neq 0 \pmod{\lambda}$, a contradiction. We conclude that $\alpha_{d'} = 0 \pmod{\lambda}$ unless $p|d'$; so g is congruent $\pmod{\lambda}$ to a form in the image of B_p, which implies that $a_n(g) = 0 \pmod{\lambda}$ unless $p|n$. This in turn implies that $\tilde{g}$ is fixed by the action not only of the normalizer of nonsplit Cartan in $\mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})$, but of a Borel subgroup as well. So $\tilde{g}$ is a modular form on $\Gamma_0(d)$ and is therefore killed by α, a contradiction. □

If f has the Fourier expansion $\sum a_n q^n$, write $f \otimes \chi$ for the modular form $\sum \chi(n) a_n q^n$. Then $f \otimes \chi$ is a newform of some level N, and in particular there is an associated quotient $A_{f\otimes\chi}$ of $J_0(N)$. Moreover, the abelian varieties $A'_f \otimes \chi$ and $A_{f\otimes\chi}$ are isogenous over $\mathbb{Q}$.

Proposition 3.9. *Suppose K is an imaginary quadratic field, and $\chi = \chi_K$ is the associated Dirichlet character. For all sufficiently large p, there exists a weight 2 cusp form f, which is either*

- *a newform in $S_2(\Gamma_0(dp^2))$ with $w_p f = f$ and $w_d f = -f$;*
- *a newform in $S_2(\Gamma_0(d'p^2))$ with d' a proper divisor of d and $w_p f = f$,*

such that $A_{f\otimes\chi}(\mathbb{Q})$ is a finite group.

We first explain how to finish the proof of Proposition 3.6 assuming the result of Proposition 3.9.

Suppose E/K is a $\mathbb{Q}$-curve of degree d, meeting the hypotheses of Proposition 3.6.

First of all, suppose λ is a prime of K dividing p. If the reduction of E at λ is potentially multiplicative, then the image of the decomposition group G_λ under $\mathbb{P}\bar{\rho}_{E,p}$ lies in a Borel subgroup. On the other hand, by hypothesis this image lies in the normalizer of a non-split Cartan subgroup. We conclude that the size of this image has order at most 2, which means that K_λ contains $\mathbb{Q}(\zeta_p + \zeta_p^{-1})$. This is impossible once $p \geq 7$.

Now suppose that E has potentially multiplicative reduction over a prime ℓ not dividing p.

The cusps of $X_0^{ns}(d;p)$ have minimal field of definition $\mathbb{Q}(\zeta_p+\zeta_p^{-1})$ [11, §5], and K is linearly disjoint from $\mathbb{Q}(\zeta_p+\zeta_p^{-1})$; it follows that the cusps of $X_0^{ns}(d;p)$ which lie over $\infty \in X_0(d)$ form a single orbit under $\mathrm{Gal}(\bar{K}/K)$. If λ is a prime of $K(\zeta_p+\zeta_p^{-1})$ over ℓ, then the point $P \in X_0^{ns}(d;p)(K)$ parametrizing E reduces mod λ to some cusp c. By applying Atkin-Lehner involutions at the primes dividing d, we can ensure that P reduces to a cusp which lies over ∞ in $X_0(d)$. By the transitivity of the Galois action, we can choose λ so that P actually reduces to the cusp ∞ mod λ. Note that in order for a K-point of $X_0^{ns}(d;p)$ to reduce to ∞, the residue field $\mathcal{O}_K/\lambda$ must contain $\zeta_p+\zeta_p^{-1}$; this implies that $\ell^4 \equiv 1 \bmod p$, and in particular $\ell \neq 2,3$ when $p \geq 7$. Moreover, if p is large enough, we have $(d,\ell)=1$.

Now take f to be a form satisfying the conditions specified in Proposition 3.9. We have defined above a map

$$X_0^{ns}(d;p)^K/R \to A_f' \otimes \chi,$$

which is a formal immersion at $\bar{\infty}$ by Lemma 3.8. Apply Proposition 3.1 with $X = X_0^{ns}(d;p)/R_\lambda$, $x = \infty$, $y = P$, and $\phi = \pi_f' \circ h$. We first apply the argument of [11, Lemma 8.3] to show that the point $\phi(P)$ is torsion in $A_f'(L)$, where $L = K(\zeta_p+\zeta_p^{-1})$. Let n be an integer which kills the subgroup of $J_0^{ns}(d;p)$ generated by cusps; such an n exists by the Drinfel'd-Manin theorem. Let σ be a generator for $\mathrm{Gal}(L/K)$; then $(\sigma-1)h(P)$ is killed by n, so $nh(P)$ lies in $J_0^{ns}(d;p)(K)$. Let τ be an element of $\mathrm{Gal}(L/\mathbb{Q})$ not lying in $\mathrm{Gal}(L/K)$. Then $P^\tau = w_d P$, and

$$n\phi(P)^\tau = n[P^\tau] - n[\infty^\tau] = n[w_d P] - n[w_d \infty] = n w_d \phi(P) = -n\phi(P).$$

So $n\phi(P)$ lies in the subgroup of $A_f'(K)$ on which τ acts as -1. By the hypothesis on f, this subgroup is finite. We conclude that $\phi(P)$ is torsion.

Since $\ell > 3$, the absolute ramification index of R_λ at ℓ is at most 2. It now follows from Proposition 3.1 that y and x reduce to distinct points of X, contradicting our hypothesis on E.

It now remains only to prove Proposition 3.9.

We say a form f is *p-new* if it is not in the space of old forms arising from $S_2(\Gamma_0(dp))$.

By a theorem of Kolyvagin and Logachev [20], building on results of Gross-Zagier, Bump-Friedberg-Hofstein, and Murty-Murty, it suffices to show that there exists a weight 2 newform f on level p^2 such that $w_p f = f$ and $L(f\otimes\chi,1) \neq 0$.

Our method will be to show the stronger statement that the (suitably weighted) *average* value of the above L-function over a certain class of forms is nonzero.

Let $\mathcal{F}$ be a Petersson-orthogonal basis for $S_2(\Gamma_0(p^2))$ such that each $f \in \mathcal{F}$ is an eigenform for all Hecke operators T_ℓ where ℓ is prime to p, and for the Atkin-Lehner involution w_p.

We define an average

$$V(p) = \sum_{\substack{f\in\mathcal{F}\\ w_p f = f}} a_1(f) L(f\otimes\chi,1).$$

First of all, note that if f is a form with $w_p f = -f$, then the functional equation of $L(f\otimes\chi,s)$ has sign $\chi(-1)=-1$ by [21, Th. 6]. So in this case $L(f\otimes\chi,1)=0$. It follows that

$$V(p) = \sum_{f\in\mathcal{F}} a_1(f) L(f\otimes\chi,1).$$

We may think of a Fourier coefficient a_n as a linear functional in $\mathrm{Hom}(S_2(\Gamma_0(p^2)),\mathbb{C})$. Likewise, write L_χ for the functional sending f to $L(f\otimes\chi,1)$. Now the Petersson inner product on $S_2(\Gamma_0(p^2))$ defines an inner product on the dual space $\mathrm{Hom}(S_2(\Gamma_0(p^2)),\mathbb{C})$, and the average we are studying is

$$V(p) = (a_1, L_\chi)$$

with respect to this dual inner product.

In general, if ℓ_1, ℓ_2 are linear functionals on the space of cuspforms for some $\Gamma_0(N)$, we write $(\ell_1, \ell_2)_N$ for the corresponding Petersson product. If V is a subspace of $S_2(\Gamma_0(N))$, we write $(\ell_1, \ell_2)_V$ for the Petersson product restricted to V. Finally, if $M|N$, we write $(\ell_1, \ell_2)_N^M$ for the contribution to $(\ell_1, \ell_2)_N$ of those forms which are new on level M.

The value of $V(p)$ can be estimated using the Petersson formula, as in [13]. In particular, we will show that, for p large enough, $V(p)$ is nonzero, and thus that $L(f \otimes \chi, 1)$ is nonzero for some f in $\mathcal{F}$.

Lemma 3.10. $V(p) = 4\pi + O(p^{-2+\epsilon})$.

Proof. Immediate from the main theorem of [14]; to be precise, we have in general that

$$(a_m, L_\chi)_N = 4\pi\chi(m) + O(N^{-1+\epsilon}) \tag{3.3}$$

with constants depending only on m, χ, and ϵ. □

But Lemma 3.10 is not enough for us, since we require that there be a *p-new* form f with $L(f \otimes \chi, 1) \neq 0$. We must therefore show that the contribution of the p-old forms to $V(p)$ is close to 0. We give below a general bound for the contribution of p-old forms to $(a_m, L_\chi)_{p^2}$. This will require an argument somewhat more intricate, but no deeper, than the Petersson estimate for $V(p)$. The problem of bounding the contribution of oldforms is treated in [18], but only in case the level is square-free. We have recently learned that a paper of Akbary [1] also bounds the contribution of oldforms in a similar situation.

Remark 3.11. We expect that, for arbitrary fixed N, the contribution of p-old forms to $(a_m, L_\chi)_{Np}$ is $o((a_m, L_\chi)_{Np})$; proving this, when a high power of p divides N, seems rather complicated.

The space of p-old forms on $\Gamma_0(p^2)$ is orthogonal to the space of p-new forms. So we can decompose the inner product (a_m, L_χ) as

$$(a_m, L_\chi) = (a_m, L_\chi)^{p-new} + (a_m, L_\chi)_{p^2}^p$$

We will show that $(a_m, L_\chi)_{p^2}^p$ approaches 0 as p grows.

Lemma 3.12. *Let f be a weight 2 newform on $\Gamma_0(p)$, let $\lambda_p(f)$ be the eigenvalue of W_p on f, and let V_f be the space of forms on $\Gamma_0(p^2)$ arising from f. Let m be a positive integer prime to p. Then*

$$(a_m, L_\chi)_{V_f} = \frac{p}{p^2-1}[1 + p^{-1}\chi(p)\lambda_p(f)](f, f)^{-1} a_m(f) L_\chi(f).$$

Proof. The space V_f is spanned by $B_1 f$ and $pB_p f$.

Then

$$(a_m, L_\chi)_{V_f} = \begin{bmatrix} a_m(B_1 f) & a_m(pB_p f) \end{bmatrix} A^{-1} \begin{bmatrix} L_\chi(B_1 f) \\ pL_\chi(pB_p f) \end{bmatrix} \tag{3.4}$$

where A is the symmetric matrix defined by $A_{ij} = (p^i B_i f, p^j B_j f)$.

It follows from the definition of Petersson product that

$$A_{11} = A_{22} = [\Gamma_0(p^2) : \Gamma_0(p)](f, f) = p(f, f).$$

We will now show that

$$(B_1 f, pB_p f) = -\lambda_p(f)(f,f). \tag{3.5}$$

Recall that we can write the Petersson product of two forms f and g as

$$C \operatorname{Res}_{s=2} L(s, f \times g)$$

where C is a constant independent of f and g, and $L(s, f \times g)$ is the Rankin-Selberg L-function defined by analytic continuation of the series

$$L(s, f \times g) = \sum_{n=1}^{\infty} a_n(f) a_n(g) n^{-s}.$$

(see [5, §1.6].)

Now

$$\begin{aligned} C^{-1}(B_1 f, pB_p f) &= p \operatorname{Res}_{s=2} \sum_{n=1}^{\infty} a_n(f) a_n(B_p f) n^{-s} \\ &= p \operatorname{Res}_{s=2} \sum_{n=1}^{\infty} a_{pn}(f) a_n(f) n^{-s} p^{-s} \\ &= -p^{-1}\lambda_p(f) \operatorname{Res}_{s=2} \sum_{n=1}^{\infty} [a_n(f)]^2 n^{-s} \\ &= -p^{-1}\lambda_p(f) C^{-1}(B_1 f, B_1 f). \end{aligned}$$

We now have

$$A = (f,f) \begin{bmatrix} p & -\lambda_p(f) \\ -\lambda_p(f) & p \end{bmatrix}.$$

Note that $\lambda_p(f) = \pm 1$. So

$$A^{-1} = (f,f)^{-1}(p^2-1)^{-1} \begin{bmatrix} p & \lambda_p(f) \\ \lambda_p(f) & p \end{bmatrix}.$$

Note that $(B_p f) \otimes \chi = \chi(p) B_p(f \otimes \chi)$. Moreover, if g is any modular form,

$$L(B_p g, 1) = \int_0^{\infty} B_p g(iy) dy = \int_0^{\infty} g(ipy) dy = (1/p) L(g,1).$$

We conclude that $L_\chi(pB_p f) = \chi(p) L_\chi(f)$. So

$$\begin{aligned} (a_m, L_\chi)_{V_f} &= p(p^2-1)^{-1}(f,f)^{-1} \begin{bmatrix} a_m(B_1 f) & 0 \end{bmatrix} \begin{bmatrix} 1 & p^{-1}\lambda_p(f) \\ p^{-1}\lambda_p(f) & 1 \end{bmatrix} \begin{bmatrix} L_\chi(f) \\ \chi(p) L_\chi(f) \end{bmatrix} \\ &= p(p^2-1)^{-1}[1 + p^{-1}\chi(p)\lambda_p(f)](f,f)^{-1} a_m(f) L_\chi(f). \end{aligned}$$

□

Now $(a_m, L_\chi)^p_{p^2}$ is the sum over newforms f of level p of $(a_m, L_\chi)_{V_f}$, which by Lemma 3.12 is equal to

$$\frac{p}{p^2-1}[1 + p^{-1}\chi(p)\lambda_p(f)](f,f)^{-1} a_m(f) L_\chi(f) = \frac{p}{p^2-1}[a_m(f) - p^{-1}\chi(p) a_{mp}(f)] L_\chi(f)(f,f)^{-1}.$$

Summing the above quantity over a Petersson-orthogonal basis of newforms for $\Gamma_0(p)$ yields

$$(a_m, L_\chi)_{p^2}^p = \frac{p}{p^2-1}(a_m - p^{-1}\chi(p)a_{mp}, L_\chi)_p.$$

Now $(a_m, L_\chi)_p$ is bounded as p grows by (3.3). So it suffices to show that $(a_{mp}, L_\chi)_p$ is $o(p^2)$. We will prove a version of this fact with explicit constants, since these will be needed in the sequel.

Lemma 3.13. *Let p be a prime, m a positive integer, χ a quadratic character of conductor q prime to p. Then*

$$(a_{mp}, L_\chi)_p \leq 2\sqrt{3}m^{1/2}d(m)(1-e^{-2\pi/q\sqrt{p}})^{-1}(4\pi + 16\zeta^2(3/2)\pi^2 p^{-3/2}).$$

Proof. Let $\mathcal{F}_p$ be a Petersson-orthogonal basis of weight 2 cuspforms on $\Gamma_0(p)$. Suppose furthermore that each $f \in \mathcal{F}_p$ is a Hecke eigenform.

First of all, note that $|a_p(f)| = |a_1(f)|$, and $|a_{mp}(f)| \leq m^{1/2}d(m)|a_1(f)|$ by Weil bounds. Now $f \otimes \chi$ is a newform on level $M = pq^2$. The functional equation for $L(f \otimes \chi, s)$ tells us that, for any positive real x,

$$L(f\otimes\chi, 1) = \sum_{n>0} \chi(n)|a_n(f)|n^{-1}e^{-2\pi n/x} + \sum_{n>0}\chi(n)|a_n(w_M(f\otimes\chi))|n^{-1}e^{-2\pi nx/M}.$$

Since $f \otimes \chi$ is a newform, we have $|a_n(w_M(f\otimes\chi))| = |a_n(f)|$. We now set $x = \sqrt{M}$ and obtain the bound

$$|L(f\otimes\chi,1)| \leq 2\sum_{n>0} a_n(f)n^{-1}e^{-2\pi n/\sqrt{M}} \leq 2(\sum_{n>0} n^{-1/2}d(n)e^{-2\pi n/\sqrt{M}})|a_1(f)|.$$

The sum over n is of length approximately $\sqrt{M}$, and so has value of order at most $M^{1/4+\epsilon}$. Working the constants out is slightly intricate, so we satisfy ourselves with a much cruder bound. Since $d(n) \leq \sqrt{3n}$, we have

$$|L(f\otimes\chi,1)| \leq 2\sqrt{3}(1-e^{-2\pi/\sqrt{M}})^{-1}|a_1(f)|,$$

so

$$\begin{aligned} |(a_m, L_\chi)| &\leq \sum_{f\in\mathcal{F}} |a_m(f)||L(f\otimes\chi,1)| \\ &\leq \sum_{f\in\mathcal{F}} m^{1/2}d(m)2\sqrt{3}(1-e^{-2\pi/\sqrt{M}})^{-1}a_1(f)^2 \\ &= 2\sqrt{3}m^{1/2}d(m)(1-e^{-2\pi/q\sqrt{p}})^{-1}(a_1,a_1)_p. \end{aligned}$$

Now by Lemma 4 of [14] we have

$$|(a_1,a_1)_p| \leq 4\pi + 16\zeta^2(3/2)\pi^2 p^{-3/2}.$$

This yields the desired result. □

We have now proved that $(a_m, L_\chi)^p_{p^2}$ approaches 0 as p goes to ∞. Therefore, $(a_1, L_\chi)^{p-new}$ approaches $V(p)$ as p grows; in particular,

$$(a_1, L_\chi)^{p-new} \neq 0$$

for p sufficiently large. We have now proved Proposition 3.9, and therefore also Proposition 3.6. □

Suppose that K is a quadratic field, and E/K is a $\mathbb{Q}$-curve of degree d. If $\mathbb{P}\bar{\rho}_{E,p}$ does not surject onto $\mathrm{PGL}_2(\mathbb{F}_p)$, then the image of $\mathbb{P}\bar{\rho}_{E,p}$ is contained in a maximal subgroup of $\mathrm{PGL}_2(\mathbb{F}_p)$; that is to say, the image is contained in either a Borel subgroup, the normalizer of a Cartan subgroup, or an exceptional subgroup isomorphic to A_4, S_4, or A_5. For any given K, there are only finitely many p for which it is possible that $\mathbb{P}\bar{\rho}_{E,p}$ has image contained in an exceptional subgroup [22, Introduction]. The following theorem now follows from Propositions 3.2,3.4, and 3.6.

Theorem 3.14. *Let K be an imaginary quadratic field and d a square-free positive integer. There exists an effective constant $M_{K,d}$ such that, for all primes $p > M_{K,d}$ and all $\mathbb{Q}$-curves E/K of degree d, either*

- *the representation*
$$\mathbb{P}\bar{\rho}_{E,p} : \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{PGL}_2(\mathbb{F}_p)$$
is surjective, or
- *E has potentially good reduction at all primes not dividing 6.*

4 Solutions to $A^4 + B^2 = C^p$

We now use the results of the previous sections to study solutions to the generalized Fermat equation

$$A^4 + B^2 = C^p \tag{4.6}$$

The goal of this section is to prove the following theorem.

Theorem 4.1. *Suppose A, B, C are coprime integers such that*

$$A^4 + B^2 = C^p$$

and $p \geq 211$. Then $AB = 0$.

Suppose (A, B, C) is a solution to (4.6) which is *primitive* (i.e., $(A, B) = 1$) and which is *non-trivial* (i.e., $AB \neq 0$.) We associate to (A, B, C) a curve $E = E_{A,B,C}/\mathbb{Q}[i]$ with the Weierstrass equation

$$E_{A,B,C} : y^2 = x^3 + 2(1+i)Ax^2 + (B + iA^2)x, \tag{4.7}$$

which was first discussed by Darmon in [8] in connection with the equation $A^4 + B^4 = C^p$. We may think of E as a "generalized Frey-Hellegouarch curve" whose relationship to (4.6) is analogous to that between the usual Frey-Hellegouarch curve and Fermat's equation.

Note that if (A, B, C) is a solution to (4.6), then so is $(A, -B, C)$. We therefore can and do assume that $B \equiv 0, 2, 3 \bmod 4$.

Write σ for the non-trivial element of $\mathrm{Gal}(\mathbb{Q}[i]/\mathbb{Q})$. The map

$$\mu : (x, y) \mapsto (\frac{1}{2}i(y^2/x^2), -\frac{1}{4}(1-i)y(B + iA^2 - x^2)/x^2).$$

is a degree 2 isogeny from E to its Galois conjugate E^σ. Therefore, $E/\mathbb{Q}[i]$ is a $\mathbb{Q}$-curve of degree 2.

One computes

$$\begin{aligned} E_4(E, \omega) &= 80iA^2 - 48B \\ \Delta(E, \omega) &= -64i(A^2 + iB)(A^2 - iB)^2, \end{aligned}$$

where ω is the Weierstrass differential $dx/2y$ with respect to the Weierstrass equation (4.7). Because $(A, B) = 1$, we have that $E/\mathbb{Q}[i]$ is semistable away from 2, and has multiplicative reduction at an odd prime $\mathfrak{p}$ of $\mathbb{Q}[i]$ precisely when $\mathfrak{p}$ divides C.

Remark 4.2. Suppose that A and B are chosen so that $A^4 + B^2$ is a prime number ℓ. Then E has good reduction away from primes of $\mathbb{Q}[i]$ dividing 2 and ℓ. Moreover, the restriction of scalars $\mathrm{Res}_{\mathbb{Q}[i]/\mathbb{Q}} E$ is an abelian surface over $\mathbb{Q}$ which has good reduction away from 2 and ℓ. We know, by the theorem of Iwaniec and Friedlander [16], that there are infinitely many choices of A, B such that $A^4 + B^2$ is prime; it follows that there are infinitely many abelian surfaces over $\mathbb{Q}$ whose bad reduction is supported at 2 and a single odd prime. It is interesting that we know this fact for abelian surfaces, but not for elliptic curves over $\mathbb{Q}$!

We now embark on an analysis of the Galois representation $\mathbb{P}\bar{\rho}_{E,p}$. We will eventually show that, when p is large, this representation surjects onto $\mathrm{PGL}_2(\mathbb{F}_p)$.

We can define a lifting of $\mathbb{P}\bar{\rho}_{E,p}$ to an actual representation as follows. The abelian surface $A = \mathrm{Res}_{\mathbb{Q}[i]/\mathbb{Q}} E$ is an abelian surface with real multiplication by $\sqrt{2}$. Let $\mathfrak{p}|p$ be a prime of $\mathbb{Z}[\sqrt{2}]$. We define

$$\bar{\rho}_{E,\mathfrak{p}} : \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(\bar{\mathbb{F}}_p)$$

to be the mod p Galois representation attached to A. Note that $\mathbb{P}\bar{\rho}_{E,p}$ is, as the notation suggests, the projectivization of $\bar{\rho}_{E,\mathfrak{p}}$, and that $\bar{\rho}_{E,\mathfrak{p}}|\,\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}[i])$ is precisely the Galois representation $E[p](\bar{\mathbb{Q}})$.

We want to show, first of all, that $\mathbb{P}\bar{\rho}_{E,p}$ is irreducible. We begin with an elementary lemma on small primes.

Lemma 4.3. *There exists a prime ℓ greater than 3 which divides C.*

Proof. If A and B were both odd, then C^p would be congruent to 2 mod 4, which is not possible. So C is odd. Moreover, C cannot be divisible by 3, since $A^4 + B^2 = 0$ has no nonzero solutions over $\mathbb{F}_3$. Finally, $C \neq 1$, since (A, B, C) is a non-trivial solution to (4.6). We conclude that there exists a prime ℓ greater than 3 which divides C; it follows that E has multiplicative reduction at primes of $\mathbb{Q}[i]$ over ℓ. □

Lemma 4.4. *$\bar{\rho}_{E,\mathfrak{p}}$ is modular.*

Proof. Since 3 does not divide C, the curve E has good reduction at 3, so $\mathbb{P}\bar{\rho}_{E,p}$ is unramified at 3. The modularity of E now follows from [15, Th. 5.2]. □

Proposition 3.2 and Lemma 4.3 imply that $\bar{\rho}_{E,p}$ is irreducible. Our next goal is to compute the Serre invariants $N = N(\bar{\rho}_{E,\mathfrak{p}}), k = k(\bar{\rho}_{E,\mathfrak{p}})$, and $\epsilon = \epsilon(\bar{\rho}_{E,\mathfrak{p}})$. By Ribet's theorem and Lemma 4.4, we now have

$$\bar{\rho}_{E,\mathfrak{p}} \cong \bar{\rho}_{f,p}$$

for some f in $S_k^{new}(N,\epsilon)$. Note that $\det \bar{\rho}_{E,\mathfrak{p}}$ is cyclotomic, which implies that ϵ is trivial and $k \cong 2 \mod p-1$.

We now use the fact that $C^p = A^4 + B^2$ is a pth power. This fact implies that every odd prime ℓ dividing $\Delta(E,\omega)$ satisfies $p|\mathrm{ord}_\ell \Delta$. By the theory of the Tate curve, this implies that $\bar{\rho}_{E,\mathfrak{p}}|\,\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}[i])$ is unramified away from 2 and p, and so $\bar{\rho}_{E,\mathfrak{p}}$ is unramified away from 2 and p. So N is a power of 2.

The fact that $p|\mathrm{ord}_{\mathfrak{q}}\Delta$ for any prime $\mathfrak{q}|p$ of $\mathbb{Z}[i]$ means, again using the Tate curve, that $E[p]/\mathbb{Q}[i]$ extends to a finite flat group scheme $\mathcal{G}$ over the completion $\mathbb{Z}[i]_{\mathfrak{q}}$. Since $\mathbb{Z}[i]/\mathbb{Z}$ is unramified at p, this extension is unique. By the étaleness of $\mathbb{Z}[i]/\mathbb{Z}$ at p, we can descend $\mathcal{G}\oplus\mathcal{G}$ to a a finite flat group scheme over $\mathbb{Z}_p$ extending $A[p]/\mathbb{Q}$. This means that $\bar{\rho}_{E,\mathfrak{p}}$ is *finite* in the sense of [27, 2.8], and so $k=2$ by [27, Prop. 4].

It remains to pin down N precisely, which we accomplish by means of Tate's algorithm.

Proposition 4.5. *$N = 32$ if A is even and 256 if A is odd.*

Proof. We begin by using Tate's algorithm to compute the local conductor f of the elliptic curve $E/\mathbb{Q}[i]$ at the prime $\pi = 1+i$. Recall that f is the multiplicity of the ideal π in the conductor of E. We refer to [28, IV. §9] for all facts about Tate's algorithm.

First of all, we will translate x by 1, which gives us a new Weierstrass equation

$$y^2 = x^3 + (3+2(1+i)A)x^2 + (3+4(1+i)A + B + iA^2)x + 2(1+i)A + 1 + B + iA^2.$$

With respect to this equation, $b_2 = (8+8i)A + 12$. If A is odd and B even, then π^2 does not divide a_6, so E has reduction type II and $f = \mathrm{ord}_\pi(\Delta) = 12$. (See [28, IV. §9,Step 3].)

If A is even and B odd, we define new variables by $y' = y - x$ and $x' = x - 1 - i$. This change of variables gives rise to the Weierstrass equation

$$\begin{aligned}(y')^2 + 2x'y' + 2(1+i)y' = \\ (x')^3 + (5+3i+2(1+i)A)(x')^2 + ((B+iA^2) + (4+12i)A + 7 + 10i)x' \\ +(B+iA^2)(i+2) + (-2+14i)A + 9i + 2.\end{aligned}$$

We are now in the situation of [28, IV. §9,Step 7], so E has reduction of type I_n^* for some n. Note that $\pi^3|a_3$. Also, π^5 divides $(-2+14i)A$ and $8i$, so we have $a_6 \cong (B+iA^2+1)(i+2) \mod \pi^5$. Recall that we've assumed B is not congruent to 1 mod 4. So $B - 2A \equiv 3 \mod 4$.

If $B-2A$ is congruent to 7 mod 8, we see that $\pi^5|a_6$. So the polynomial $Y^2 + \pi^{-2}a_3 Y - \pi^{-4}a_6 Y$ has a double root over $\mathbb{F}_2$ at $Y=0$. Moreover, in this case

$$a_4 \cong B + 7 + 10i \cong B - 1 + 2i$$

mod π^4. So $\mathrm{ord}_\pi a_4 = 3$, which implies that the polynomial $\pi^{-1}a_2X^2 + \pi^{-3}a_4X + \pi^{-5}a_6$ has distinct roots in $\mathbb{F}_2$. We conclude in this case that E has reduction type I_2^* and $f = \mathrm{ord}_\pi(\Delta) - 6 = 6$.

Suppose on the other hand that $B-2A$ is congruent to 3 mod 8. Then we change variables by setting $y'' = y' + 2$. This change of variables causes a_6 to become a multiple of π^5, while the valuations of a_4 and a_2 do not change. So once again we are in the situation of reduction type I_2^* and $f = 6$.

The quantity f computed above is the Artin conductor of T_pE considered as a p-adic Galois representation of $\mathrm{Gal}(\bar{\mathbb{Q}}_2/\mathbb{Q}_2[i])$. Let ρ_* be the the 4-dimensional representation of $\mathrm{Gal}(\bar{\mathbb{Q}}_2/\mathbb{Q}_2)$ induced from T_pE. Then $\rho_* \cong T_pA$, where A is the restriction of scalars of E described above. We have from [24, §1] that $f(\rho_*) = f + 2\dim T_pE = f+4$. So $f(\rho_*)$ is either 10 or 16.

By examination of the j invariant, we see that E has potentially good reduction at π. It follows that the inertia group $I_2 \subset \mathrm{Gal}(\bar{\mathbb{Q}}_2/\mathbb{Q}_2)$ acts on T_pA through a finite quotient G, whose order is not divisible by any prime greater than 3. Let $\rho_{E,\mathfrak{p}}$ be the 2-dimensional representation of $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ on $T_{\mathfrak{p}}A$, and let $f_{\mathfrak{p}}$ be the conductor of $\rho_{E,\mathfrak{p}}|\,\mathrm{Gal}(\bar{\mathbb{Q}}_2/\mathbb{Q}_2)$. If $\mathfrak{p} = p$ is inert, it is immediate that $f(\rho_*) = 2f_{\mathfrak{p}}$. If, on the other hand, $\mathfrak{p}$ and $\mathfrak{p}'$ are split primes of $\mathbb{Q}(\sqrt{2})$ lying over p, then $f_{\mathfrak{p}}$ and $f_{\mathfrak{p}'}$ are both equal to the 2-part of the conductor of the modular abelian variety A, and again we get

$$f(\rho_*) = f_{\mathfrak{p}} + f_{\mathfrak{p}'} = 2f_{\mathfrak{p}}.$$

We conclude that $f_{\mathfrak{p}}$ is either 5 or 8. Moreover, the fact that $|G|$ is prime to p implies that the 2-part of the conductor of $\bar{\rho}_{E,\mathfrak{p}}$ is identical with $f_{\mathfrak{p}}$. This completes the proof. □

We have now established that $\bar{\rho}_{E,\mathfrak{p}}$ is isomorphic to $\bar{\rho}_{f,p}$, where f is a weight 2 newform of level 32 or 256. In fact, the newforms of these levels are all associated to elliptic curves (not necessarily defined over $\mathbb{Q}$) with complex multiplication by $\mathbb{Q}[i]$ or $\mathbb{Q}[\sqrt{-2}]$. In particular, the image of $\mathbb{P}\bar{\rho}_{E,p}$ is the normalizer of a Cartan subgroup in $PGL_2(\mathbb{F}_p)$.

Suppose the image of $\mathbb{P}\bar{\rho}_{E,p}$ lies in the normalizer of a split Cartan subgroup. Then it follows from Proposition 3.4 that E has good reduction away from 6. But this contradicts Lemma 4.3.

We conclude that the image of $\mathbb{P}\bar{\rho}_{E,p}$ must be the normalizer of a non-split Cartan subgroup. We now use this fact to bound p.

Proposition 4.6. *Let $E_0/\mathbb{Q}[i]$ be a $\mathbb{Q}$-curve of degree 2, and suppose $\mathbb{P}\bar{\rho}_{E,p}$ has image contained in the normalizer of a non-split Cartan subgroup of $\mathrm{PGL}_2(\mathbb{F}_p)$, for some $p \geq 211$. Then E_0 has potentially good reduction for all primes of $\mathbb{Q}[i]$.*

Proposition 4.6 completes the proof of Theorem 4.1. For we have shown that if (A, B, C) is a solution to (4.6), and $E = E_{A,B,C}$, then $\mathbb{P}\bar{\rho}_{E,p}$ has as image the normalizer of a non-split Cartan subgroup of $\mathrm{PGL}_2(\mathbb{F}_p)$. If $p \geq 211$, then Proposition 4.6 shows that E has good reduction everywhere, contradicting Lemma 4.3.

We now proceed with the proof of Proposition 4.6.

Proof. The main tools are Proposition 3.6 and the estimate for average special values of L-functions in [14].

Proposition 3.6 tells that E_0 has potentially good reduction everywhere if p is sufficiently large. We now show that $p \geq 211$ suffices. It is clear from the proof of Proposition 3.6 that E_0 has good reduction away from $6p$ whenever there exists a newform f on $S_2(\Gamma_0(p^2))$ satisfying the conditions of Proposition 3.9. In turn, in order to prove the existence of such a form it suffices to show that the inner product

$$(a_1, L_\chi)_{p^2}^{p-new} = (a_1, L_\chi)_{p^2} - (a_1, L_\chi)_{p^2}^{p}$$

is non-zero. We have by Lemma 3.12 that

$$(a_1, L_\chi)_{p^2}^{p} = p(p^2-1)^{-1}(a_1 - p^{-1}\chi(p)a_p, L_\chi)_p.$$

We now use Theorem 1 of [14] to show that $(a_1, L_\chi)^{p-new}$ is nonzero. The theorem shows that $(a_m, L_\chi)_N$ is $4\pi\chi(m) + O(N^{-1}\log(N)d(N))$, where $d(N)$ is the number of divisors of N andthe implied constants are explicit functions of m, χ. Precisely, we obtain that, for $p \geq 211$,

$$|(a_1, L_\chi)_{p^2} - 4\pi| \leq 4.37$$

(In the applications here, we always take the parameter σ in [14] to be $8/\pi$.)

At level p, the same theorem gives

$$|(a_1, L_\chi)_p| \leq 786$$

when $p > 211$.

Finally, Lemma 3.13 shows that

$$|(a_p, L_\chi)_p| \leq 437.$$

So we find

$$|(a_1, L_\chi)^{p-new}| \geq 4\pi - 4.37 - \frac{211}{211^2 - 1}(786 + 437/211) > 4.$$

This proves Proposition 4.6. □

For p smaller than 211, the argument above shows that $A^4 + B^2 = C^p$ has no nontrivial solutions if we can prove the existence of a modular form satisfying the conditions of Proposition 3.9. To be precise, we have shown

Proposition 4.7. *Let $p > 13$ be prime, and suppose there exists either*

- *a newform in $S_2(\Gamma_0(2p^2))$ with $w_p f = f$ and $w_2 f = -f$; or*
- *a newform in $S_2(\Gamma_0(p^2))$ with $w_p f = f$,*

such that $L(f \otimes \chi, 1) \neq 0$. Then the equation $A^4 + B^2 = C^p$ has no primitive non-trivial solutions.

Verification of the existence of such a modular form is, in principle, a finite computation. In practice, it is beyond the reach of current computers to compute the Fourier coefficients of a newform of level $2p^2$ when p is as large as 100. It seems probable that by exploiting various tricks and carrying out more complicated computations, we will be able to show that Proposition 4.7 applies for all primes p between 17 and 211. We will discuss this problem in a later paper.

References

[1] A. Akbary. Non-vanishing of weight k modular L-functions with large level. *J. Ramanujan Math. Soc.*, 14(1):37–54, 1999.

[2] A. O. L. Atkin and W. C. W. Li. Twists of newforms and pseudo-eigenvalues of W-operators. *Invent. Math.*, 48(3):221–243, 1978.

[3] A.O.L. Atkin and J. Lehner. Hecke operators on $\Gamma_0(m)$. *Math. Ann.*, 185:134–160, 1970.

[4] N. Bruin. The Diophantine equations $x^2 \pm y^4 = \pm z^6$ and $x^2 + y^8 = z^3$. *Compositio Math.* 118 (1999), no. 3, 305–321.

[5] D. Bump. *Automorphic forms and representations*. Cambridge University Press, 1997.

[6] I. Chen. On relations between Jacobians of certain modular curves. *J. Algebra*, 231(1):414–448, 2000.

[7] H. Darmon. The equations $x^n + y^n = z^2$ and $x^n + y^n = z^3$. *Internat. Math. Res. Notices*, 10:263–274, 1993.

[8] H. Darmon. Serre's conjectures. In V. Kumar Murty, editor, *Seminar on Fermat's Last Theorem*, number 17 in CMS Conference Proceedings, pages 135–153, 1995.

[9] H. Darmon. Rigid local systems, Hilbert modular forms, and Fermat's last theorem. *Duke Math. J.*, 102(3):413–449, 2000.

[10] H. Darmon and A. Granville. On the equations $z^m = F(x, y)$ and $Ax^p + By^q = Cz^r$. *Bull. London Math. Soc.*, 27(6):513–543, 1995.

[11] H. Darmon and L. Merel. Winding quotients and some variants of Fermat's last theorem. *J. Reine Angew. Math.*, 490:81–100, 1997.

[12] B. DeSmit and B. Edixhoven. Sur un résultat d'Imin Chen. *Math. Res. Lett.*, 7(2–3):147–153, 2000.

[13] W. Duke. The critical order of vanishing of automorphic L-functions with large level. *Invent. Math.*, 119(1):165–174, 1995.

[14] J. Ellenberg. On the error term in Duke's estimate for the average special value of L-functions. To appear, *Canad. Math. Bull.*

[15] J. Ellenberg and C. Skinner. On the modularity of $\mathbf{Q}$-curves. *Duke Math. J.*, 109(1):97–122, 2001.

[16] J. Friedlander and H. Iwaniec. The polynomial $X^2 + Y^4$ captures its primes. *Ann. of Math (2)* 148(3): 945–1040, 1998.

[17] A. Grothendieck. Elements de geometrie algebrique, IV_4. *Publ. Math. IHES*, 32, 1967.

[18] H. Iwaniec, W. Luo, and P. Sarnak. Low lying zeros of families of L-functions. *Inst. Hautes Études Sci. Publ. Math*, 91:55–131, 2000.

[19] N. Katz and B. Mazur. *Arithmetic Moduli of Elliptic Curves*. Princeton University Press, 1985.

[20] V.A. Kolyvagin and D. Yu Logachev. Finiteness of the Shafarevich-Tate group and the group of rational points for some modular abelian varieties. *Leningrad Math. J.*, 1(5):1229–1253, 1990.

[21] W. C. W. Li. Newforms and functional equations. *Math. Ann.* 212: 285–315, 1975.

[22] B. Mazur. Modular curves and the Eisenstein ideal. *Publ. Math. I.H.E.S.*, 47:33–186, 1977.

[23] B. Mazur. Rational isogenies of prime degree. *Invent. Math.*, 44(2):129–162, 1978.

[24] J.S. Milne. On the arithmetic of abelian varieties. *Invent. Math.*, 17:177–190, 1972.

[25] F. Momose. Rational points on the modular curves $X_{split}(p)$. *Compositio Math.*, 52(1):115–137, 1984.

[26] K. Ribet. On the equation $a^p + 2b^p + c^p = 0$. *Acta Arith.*, 79(1):7–16, 1997.

[27] J.P. Serre. Sur les représentations modulaires de degré 2 de $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$. *Duke Math. Jour.*, 54(1):179–230, 1987.

[28] J. Silverman. *Advanced topics in the arithmetic of elliptic curves*. Springer-Verlag, 1994.